In this document, we describe the different use cases for CXL 2.0 Fabric Manager. We provide the mechanism and workflows for integrating CXL Fabric Manager with the [Open Fabric Management Framework](https://www.openfabrics.org/openfabrics-management-framework/) (OFMF) developed by Open Fabrics Alliance.

Based on recommendations from the CXL consortium, we focus on the target use cases for fabric-attached memory (FAM) pooling in CXL 2.0: (1) Single logical device where a FAM node’s resources consist of a single memory region bound to a host, (2) Multi logical device (MLD) where FAM node resources may be partitioned into multiple memory regions, each of which is bound to a different host.

Shape

Description automatically generated

Below, we discuss the management operations carried out by Fabric Manager (FM) to enable the above use cases.

1. **Initial resource discovery workflow**

Diagram

Description automatically generated

|  |  |
| --- | --- |
| Use Case Description | Initial discovery of CXL switch and attached resources by CXL Fabric Manager |
| Actors | CXL Fabric Manager (FM), CXL Provider and OFMF |
| Description | * FM discovers the connected switch and FAM node, and then initiates management commands to discover capabilities of the attached devices * FM propagates the discovered resource information to Provider and OFMF service. * After discovery, FM is ready for event notifications (e.g., hot removal of FAM node) and can enable host to FAM node binding. |
| Comments | Once the system components including switch and FAM node are powered up, they can be discovered irrespective of the host power status. BMC and the host can have a management link using the MCTP or Redfish host interface. Here, we focus on the discovery of capabilities of attached FAM node. and prepared for host binding to FAM node |
| Input Data | MCTP endpoint information for connected physical resources including the CXL switch and FAM node. FM initiates the discovery workflow for each endpoint. |
| Preconditions | * BMC, switch and FAM node are in-service with active management network links (e.g., on SMBus, PCIe VDM at the physical layer) * CXL switch is released from reset and loads its initial configuration from non-volatile memory. Ports are released from reset to link up. * FM is up and communicating with devices using MCTP. Switch and FAM nodes implement FM APIs * FM has sent event notification through its Provider to the OFMF Provider Manager, so that OFMF is actively managing the FM. |
| Postconditions | * Provider data store contains the newly discovered resource information   + CXL provider has queried the FM data store for new fabric state information in its next polling cycle.   + CXL provider has translated the queried fabric state information, from CXL native model to standard Redfish entities.   + CXL provider has sent event notification to OFMF Resource Inventory service about the updated fabric state. * OFMF Redfish tree contains the description of the CXL switch, FAM node and associated properties (e.g., number of ports on switch, memory capacity on FAM node) * Fabric hardware state matches the state information in Provider and OFMF service data stores |
| Trigger | Once the switch and FAM node are connected over the management link, device enumeration on the physical bus is performed by the MCTP discovery protocol. Completion of MCTP discovery triggers the FM discovery workflow |
| Normal Flow | FM detects the new resources and their property information, and this is percolated to OFMF services for querying by OFMF clients   * Initial MCTP discovery protocol detects switch and FAM node endpoints * FM issues *Get Supported Logs* command to get device specific logs (identified by UUID) on the switch and FAM node. * For each log ID, FM issues *Get Command Effect Log* command which outputs the command opcode along with description of the command effects for the opcode. * A switch is detected, so FM issues *Identify Switch Device* command to get a response of switch capabilities and capacity configuration, including physical port count, port ID of the management port on the switch, number of virtual CXL switches, etc. * FM queries individual port information on the switch by issuing *Get Physical Port State* with the number of ports requested and port ID list. Response includes port information including whether it’s an upstream port (connected to host) or downstream port (connected to devices), along with connected device CXL version, type and number of supported logical devices. * For each host port, FM issues *Get Virtual CXL Switch Info* with the number of virtual CXL switches (VCS) for which information is requested, along with their IDs. Response includes number of VCSs returned, whether each VCS is enabled/disabled, host binding status, etc. * Since a FAM node is detected, FM checks if it is a MLD by issuing *MLD Component Command Set*, to get number of supported logical devices, per-logical device information including memory size, allocation region, etc. * FM finally sets the event notification policy. Get Event Interrupt Policy enables checking the current event notification policy and it can be modified using Set Event Interrupt Policy * FM updates the newly discovered switch and FAM node in its data store. The data store uses native fabric model based on Redfish CXL extensions. |

1. **[WIP] Binding host to LD on a switch**

Chart

Description automatically generated

|  |  |
| --- | --- |
| Use Case Description | FM configures CXL switch to bind a host to memory region in FAM node |
| Actors | FM, CXL 2.0 switch, FAM node, host |
| Description | * After initial resource discovery, CXL FM receives a request to bind a host connected to the switch, with memory region in FAM node attached on the same switch. * FM configures the virtual CXL switch ports to the downstream physical ports, for binding to the FAM node * If MLD device, CXL FM configures the logical device binding |
| Comments |  |
| Input Data | FM receives request from CXL provider for the host and LD to bind. Input includes:   * Physical port where the device is attached * Virtual CXL switch (VCS) ID, mapped 1:1 to each host * Virtual PPB index within VCS * In case of MLD, LD ID for the memory chunk |
| Preconditions | * BMC, switch, host and FAM node are in-service with active management network links (e.g., SMBus, PCIe) * Initial resource discovery is complete * The host and FAM memory region are unbound (if binding already exists, OFMF service can handle the client request) |
| Postconditions | * CXL switch configured with binding for host to logical device on FAM node * Binding information percolated to the OFMF service layer |
| Trigger | Binding request from CXL provider |
| Normal flow | TBD |

1. **[WIP] Unbinding host and LD on a switch** Chart

   Description automatically generated
2. **[WIP] Hot add of device**

Chart

Description automatically generated

1. **[WIP] Managed hot removal from an unbound port**

Chart

Description automatically generated